

**PENGEMBANGAN INSTRUMEN PENILAIAN *THREE-TIER*
MULTIPLE CHOICE (Three-TMC) UNTUK MENGUKUR
KETERAMPILAN BERPIKIR KRITIS SISWA PADA
MATERI KIMIA LARUTAN PENYANGGA
KELAS XI SMA**

Wiad Rosyana¹, Ashadi² dan Sri Mulyani³

¹Program Studi Magister Pendidikan Sains FKIP Universitas Sebelas Maret
Surakarta, 57126, Indonesia
wiadrosyana@gmail.com

²Program Studi Magister Pendidikan Sains FKIP Universitas Sebelas Maret
Surakarta, 57126, Indonesia
ashadiuns2014@gmail.com

³Program Studi Magister Pendidikan Sains FKIP Universitas Sebelas Maret
Surakarta, 57126, Indonesia
srimulyaniuns@staff.uns.ac.id

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen penilaian *Three-Tier MC* dengan kriteria baik ditinjau dari validitas, reliabilitas, analisis kualitas butir soal, dan analisis pengecoh serta mengembangkan instrumen penilaian *Three-Tier MC* menjadi instrumen penilaian yang dapat dijadikan sebagai alat ukur keterampilan berpikir kritis siswa pada materi larutan penyangga. Penelitian pengembangan instrumen penilaian *Three-Tier MC* ini menggunakan prosedur Borg and Gall. Validasi instrumen penilaian *Three-Tier MC* ini menggunakan rumus Aiken dengan 8 validator. Teknik penskoran yang digunakan merupakan penskoran yang bersifat politomi. Hasil penelitian menunjukkan bahwa instrumen penilaian *Three-Tier MC* yang telah dikembangkan termasuk dalam kategori yang baik ditinjau dari segi validitasnya yang lebih dari 0,75 dan dikatakan valid; reliabilitasnya termasuk dalam sedang, daya pembedanya tidak ada yang bernilai negatif dengan persentase jelek 8,89%, cukup baik 20%, baik 24,44%, sangat baik 46,67%; soal yang memiliki tingkat kesukaran mudah dan sukar tidak mendominasi dengan persentase soal mudah 15,56%, sedang 73,33%, sukar 11,11%; dan semua pengecoh berfungsi dengan baik yang ditunjukkan tidak adanya nilai indeks pengecoh 0,000; serta instrumen penilaian *Three-Tier MC* yang telah dikembangkan pada materi larutan penyangga dapat digunakan untuk mengukur keterampilan berpikir kritis siswa.

Kata kunci : instrumen penilaian, *Three-Tier MC*, keterampilan berpikir kritis, larutan penyangga.

Pendahuluan

Pendidikan saat ini merupakan salah satu hal yang menjadi perhatian serius bangsa Indonesia mengingat pentingnya peranan pendidikan dalam rangka menyiapkan sumber daya manusia (SDM) untuk pembangunan dan persaingan global. Selain itu, teknologi dan ilmu pengetahuan selalu mengalami

perkembangan. Oleh karena itu pemerintah berupaya untuk meningkatkan kualitas pendidikan mencakup berbagai bidang diantaranya peningkatan sarana dan prasarana, perubahan kurikulum dan proses belajar mengajar, peningkatan kualitas guru, dan usaha-usaha lain yang tercakup dalam komponen pendidikan.

Kimia merupakan salah satu bidang ilmu yang mengalami perkembangan dari waktu ke waktu. Pada jenjang Sekolah Menengah Atas (SMA), mata pelajaran kimia merupakan salah satu mata pelajaran wajib bagi siswa jurusan Ilmu Pengetahuan Alam (IPA). Materi pembelajaran kimia dianggap masih sulit dipelajari dan mata pelajaran kimia belum menjadi pilihan utama dalam belajar karena menurut mereka materi pelajaran kimia kurang menarik. Banyak teori yang sulit untuk dipahami hanya dengan satu kali penyampaian. Hal ini dapat menyebabkan prestasi belajar kimia rendah atau belum mencapai target yang diinginkan.

Larutan penyangga merupakan larutan yang dapat mempertahankan pH dengan penambahan sedikit asam atau basa. Suatu hal yang penting untuk dapat memahami materi larutan penyangga karena sangat berhubungan erat dengan kehidupan manusia. Konsep pada materi larutan penyangga merupakan konsep yang bersifat kompleks dan hierarkis. Materi larutan penyangga banyak berhubungan dengan konsep materi kimia yang lain yaitu larutan asam basa, konsep pH, persamaan reaksi, kesetimbangan, dan stoikiometri larutan.

Hasil belajar bukan merupakan satu-satunya hal yang harus diperhatikan. Berdasarkan Permendiknas Nomor 23 Tahun 2006 tentang Standar Kompetensi Lulusan Satuan Pendidikan (SKL-SP) hal yang juga perlu diperhatikan adalah bahwa siswa dapat membangun dan menerapkan informasi dan pengetahuan secara logis, kritis, kreatif, dan inovatif; menunjukkan kemampuan berpikir logis, kritis, kreatif, dan inovatif dalam pengambilan keputusan. Keterampilan berpikir kritis perlu dikembangkan sebagai persiapan siswa agar mampu menghadapi masalah kompleks sebagai akibat dari perkembangan ilmu dan teknologi. Berdasarkan "21st Century Partnership Learning Framework", terdapat beberapa kompetensi yang harus

dimiliki oleh SDM pada abad XXI, salah satunya yaitu keterampilan berpikir kritis dan memecahkan masalah (*Critical-Thinking and Problem-Solving Skills*), yaitu mampu berpikir secara kritis, lateral, dan sistemik, terutama dalam konteks pemecahan masalah (BSNP, 2013).

Berpikir kritis merupakan bagian dari alur berpikir ilmiah yaitu kemampuan berpikir yang didasarkan pada logika, kejelasan dan relevansi. Dewey berpendapat bahwa berpikir kritis merupakan berpikir reflektif yaitu "pertimbangan yang aktif, *persistent* (terus menerus), dan teliti mengenai sebuah keyakinan atau bentuk pengetahuan yang diterima begitu saja dipandang dari sudut alasan-alasan yang mendukungnya dan kesimpulan-kesimpulan lanjutan yang menjadi kecenderungannya" (Fisher, 2009). Berpikir kritis merupakan proses aktif siswa mengenai sebuah pengetahuan yang bisa diterima. Keterampilan berpikir kritis akan mempengaruhi bagaimana cara berpikir siswa dalam menerima pengetahuan. Bagaimana mereka memahami, mencari tahu lebih dalam, mengevaluasi penguasaan pengetahuan dan mengubah pemikiran terhadap pengetahuan itu sendiri.

Siswa yang memiliki keterampilan berpikir kritis tidak akan serta merta pasrah menerima apa yang disampaikan guru. Peningkatan keterampilan berpikir kritis dimaksudkan agar siswa lebih mampu berpikir secara mendalam sehingga akan dapat lebih mudah memahami konsep-konsep ilmiah. Hal ini juga sesuai dengan pendapat dari Johnson dan Siegel bahwa dengan memiliki keterampilan berpikir kritis, siswa dapat mengkonstruksi pengetahuan yang diterimanya dalam penguasaan konsep secara utuh (Prasetyowati & Suyatno, 2016).

Dalam jurnalnya, Qing menjelaskan bahwa dalam proses belajar diperlukan kemampuan berpikir sebagai

tujuan utama dalam pendidikan (Qing, Jing, Yazhuan, Ting, & Junping, 2010). Apabila siswa memiliki keterampilan berpikir kritis, mereka akan mampu mengimplementasikan teori yang diperolehnya. Miri menyadari pentingnya meningkatkan keterampilan berpikir kritis untuk menghadapi perkembangan dalam pendidikan sains. Miri melakukan penelitian dengan hasil yang menunjukkan apabila guru dengan sengaja secara terus menerus melatih siswa dengan menggunakan strategi pembelajaran untuk meningkatkan berpikir tingkat tinggi ada kesempatan yang baik untuk mengembangkan keterampilan berpikir kritis (Miri, David, & Uri, 2007).

Sesuai dengan karakteristiknya, berpikir kritis memerlukan latihan yang salah satunya dengan membiasakan mengerjakan soal-soal yang mengembangkan keterampilan berpikir kritis (Kartimi & Liliyasi, 2012). Perlu adanya alat ukur untuk mengetahui tingkat keberhasilannya. Keterampilan berpikir kritis dapat diukur dengan menggunakan instrumen yang khusus diperuntukkan untuk mengukur keterampilan berpikir kritis dengan berorientasi pada aspek-aspek yang ada di dalamnya.

Sulfivan dan Dawson mencoba untuk membuat instrumen tes keterampilan berpikir kritis berdasar situasi sains (Sulfivan & Dawson, 1979). Agar dapat mengetahui kemampuan berpikir kritis siswa pada kelas matrikulasi kimia harus menggunakan instrumen yang sesuai. Dalam penyusunan instrumen tes, ada yang mudah untuk disusun seperti indikator interpretasi tetapi juga ada yang sulit seperti indikator evaluasi. Dari instrumen yang dibuat dan pembahasan hasilnya tidak semua indikator berpikir kritis dapat diukur. Kartimi telah mengembangkan alat ukur berpikir kritis pada materi termokimia berupa soal pilihan ganda bertingkat, dimana

hasilnya menunjukkan bahwa alat ukur tersebut mampu membedakan keterampilan berpikir kritis siswa (Kartimi & Liliyasi, 2012).

Instrumen tes pilihan ganda bertingkat telah banyak dikembangkan seperti *Two-Tier Test* dan *Three-Tier Test*. *Two-Tier Test* merupakan tes berupa pilihan ganda dua tingkat, dimana tingkat pertama adalah soal pilihan ganda dengan beberapa opsi jawaban dan tingkat kedua merupakan alasan jawaban dari tingkat pertama. Bayrak mengembangkan *Two-Tier Test* (tes dua tingkat) untuk mengetahui pemahaman konsep dan miskonsepsi pada materi asam basa (Bayrak, 2013). Studi kasus yang dilakukan Chandrasegaran, mengembangkan instrumen diagnosa *Two-Tier Multiple-Choice* dengan menggunakan prosedur yang diusulkan oleh Treagust (Chandrasegaran, Treagust, & Mocerino, 2007). Kemudian instrumen tes dua tingkat ini berkembang menjadi tiga tingkat yang disebut *Three-Tier Test*. Eryilmaz dan Pesman mengembangkan *Three-Tier Test* untuk mengetahui miskonsepsi siswa (Eryilmaz, 2010; Peşman & Eryilmaz, 2010). Pada *Three-Tier Test*, terdapat tingkat ketiga yang merupakan pernyataan yakin-tidak yakin dari siswa.

Instrumen tes pilihan ganda bertingkat ini kemudian dikembangkan lagi tidak hanya sebagai instrumen tes diagnosa. Salah satu pengembangan tes pilihan ganda bertingkat dilakukan oleh Yamtinah untuk mengukur keterampilan proses sains pada pelajaran kimia dengan teknik penskoran *Graded Response Model* (GRM) (Yamtinah, Saputro, & Utami, 2016). GRM adalah teknik penskoran yang sesuai dengan penskoran yang bersifat politomus. Nofiana juga mengembangkan tes pilihan ganda dua tingkat untuk mengukur keterampilan berpikir tingkat tinggi. Nofiana berpendapat bahwa kemampuan peserta didik untuk berpikir kritis dan kreatif termasuk dalam bentuk keterampilan

berpikir tingkat tinggi (Nofiana, Sajidan, & Puguh, 2014).

Di Indonesia sendiri masih jarang penggunaan instrumen penilaian untuk mengukur keterampilan berpikir kritis. Merujuk pada perkembangan bentuk instrumen *Three-Tier Test* yang dikembangkan Eryilmaz dan Pesman serta pengembangan alat ukur berpikir kritis yang dilakukan Kartimi dengan soal pilihan ganda bertingkat maka perlu dikembangkan instrumen penilaian *Three-Tier Multiple Choice (Three-Tier MC)* untuk mengukur keterampilan berpikir kritis siswa. Instrumen penilaian harus disusun sesuai dengan indikator dari keterampilan berpikir kritis. Terdapat enam indikator keterampilan berpikir kritis yaitu interpretasi, analisis, evaluasi, inferensi, penjelasan dan pengaturan diri. Keenam indikator tersebut adalah inti dari keterampilan berpikir kritis yang menurut para ahli termasuk didalam keterampilan kognitif (Facione, 2015).

Instrumen penilaian *Three-Tier MC* terdiri dari tiga tingkat soal pilihan ganda pada setiap nomornya. Pada tingkat pertama di *Three-Tier MC* ini nantinya siswa dituntut untuk bisa mengidentifikasi soal yang disajikan. Kemudian menganalisis dan menarik kesimpulan dengan memilih salah satu jawaban yang telah disajikan. Pada tingkat kedua, diberi beberapa pilihan pernyataan alasan atas jawaban dari tingkat pertama. Tingkat ini dapat mewakili analisis dan refleksi pengetahuan dari siswa. Tingkat ketiga merupakan soal yang berupa pernyataan dengan jawaban benar-salah yang berhubungan dengan pengetahuan tingkat sebelumnya. Tingkat ini mengandung konsep yang sama dengan tingkat pertama, atau dengan kata lain tingkat ketiga ini adalah pengulangan konsep dari tingkat

pertama. Sehingga instrumen penilaian ini dapat dijadikan sebagai alat ukur keterampilan berpikir kritis.

Metode Penelitian

Penelitian ini termasuk penelitian dan pengembangan dimana produk yang dikembangkan adalah instrumen penilaian *Three-Tier Multiple Choice (Three-Tier MC)* untuk mengukur keterampilan berpikir kritis siswa pada materi larutan penyangga.

Penelitian dan pengembangan ini mengacu pada model *RnD* yang dikemukakan oleh Borg and Gall. Tahapan yang dilakukan dalam penelitian dan pengembangan ini yaitu penelitian dan pengumpulan informasi; membuat perencanaan; mengembangkan rancangan awal produk (*draft*); melakukan uji coba lapangan awal; melakukan revisi produk utama; melakukan uji lapangan utama; melakukan revisi produk operasional; melakukan uji lapangan operasional; dan melakukan revisi produk akhir (Borg & Gall, 1983).

Subjek uji coba instrumen penilaian *Three-Tier MC* ini adalah siswa-siswi yang berasal dari sekolah dengan kategori kemampuan tinggi, sedang, dan rendah. Pada uji lapangan awal, produk diujicobakan pada 9 siswa setiap sekolah. Uji lapangan utama, produk diujicobakan pada 1 kelas dari tiap sekolah. Pada uji coba operasional, produk diujicobakan pada 2 kelas dari tiap sekolah.

Teknik pengumpulan data dalam penelitian ini adalah dengan metode angket dan tes. Jenis data terdiri dari data kualitatif dan kuantitatif. Data kuantitatif diperoleh dari uji validitas menggunakan rumus Aiken, dan analisis butir soal berupa reliabilitas, daya pembeda, tingkat kesukaran soal, serta indeks pengecoh menggunakan program *software ITEMAN*.

Hasil Penelitian dan Pembahasan

Tahapan awal yang dilakukan dalam penelitian dan pengembangan ini adalah melakukan studi pendahuluan. Studi pendahuluan bertujuan untuk mengetahui permasalahan yang tengah terjadi di lapangan. Kemudian mengerucutkan objek masalah pada sekolah tertentu untuk melakukan penelitian lebih lanjut. Studi pendahuluan dilakukan dengan dua cara yaitu studi literatur dan studi lapangan.

Penelitian dan pengembangan produk sebagai sarana untuk meningkatkan keterampilan berpikir kritis siswa sudah banyak. Akan tetapi belum banyak pengembangan tentang instrumen tes untuk mengukur tingkat keterampilan berpikir kritis siswa. Sulfivan dan Dawson mencoba untuk membuat instrumen tes keterampilan berpikir kritis berdasar situasi sains. Agar dapat mengetahui kemampuan berpikir kritis siswa pada kelas matrikulasi kimia harus menggunakan instrumen yang sesuai (Sulfivan & Dawson, 1979). Dalam penyusunan instrumen tes, ada yang mudah untuk disusun seperti indikator interpretasi tetapi juga ada yang sulit seperti indikator evaluasi. Dari instrumen yang dibuat dan pembahasan hasilnya tidak semua indikator berpikir kritis dapat diukur. Kartimi telah mengembangkan alat ukur berpikir kritis pada materi termokimia, dimana hasilnya menunjukkan bahwa alat ukur tersebut mampu membedakan keterampilan berpikir kritis siswa. Alat ukur yang dibuatnya berupa soal pilihan ganda bertingkat (Kartimi & Liliyasi, 2012; Kartimi et al., 2012).

Evaluasi pembelajaran yang digunakan oleh guru bervariasi. Pada sekolah kategori tinggi, bentuk soal-soal yang digunakan dalam evaluasi lebih bervariasi dibanding dua sekolah lainnya. Bentuk soalnya berupa pilihan ganda, uraian dan jawaban singkat. Untuk sekolah kategori sedang lebih sering

menggunakan soal bentuk uraian, sedangkan sekolah kategori rendah lebih sering soal pilihan ganda.

Jenjang soal pun guru dari masing-masing sekolah mempunyai pendapat masing-masing. Guru dari sekolah kategori tinggi memaparkan jenjang soal yang dibuatnya mencakup C1 sampai C6. Pembuatan soal ini selain berdasarkan tuntutan dari tujuan pembelajaran yang berorientasi pada taksonomi Bloom, juga didasarkan pada soal-soal ujian nasional. Dari analisis yang dilakukan, guru mendapati bahwa soal-soal ujian nasional sudah berorientasi pada keterampilan berpikir tingkat tinggi. Dimana siswa tidak hanya sekedar menghafal rumus-rumus, tetapi juga dituntut untuk memahami konsep-konsep. Sedangkan pada sekolah kategori sedang dan rendah, lebih sering menggunakan soal jenjang C1 sampai C4. Untuk C5 dan C6, masih jarang diterapkan karena soal jenjang tersebut tidak mudah untuk dibuat. Selain itu, tidak jarang siswa juga mengeluhkan soal sulit untuk dikerjakan.

Langkah selanjutnya adalah *planning* atau perencanaan. Setelah menganalisis temuan yang didapatkan dari studi pendahuluan, sebagai salah satu langkah yang diharapkan menjadi solusi atas permasalahan maka dikembangkanlah suatu instrumen penilaian bentuk pilihan ganda yang memiliki efisiensi waktu dalam pengerjaan maupun pengoreksian, mengetahui pemahaman siswa, mengurangi peluang siswa dalam menebak jawaban, serta mengetahui cara berpikir siswa.

Instrumen penilaian yang dikembangkan ini juga diharapkan akan memperkaya jenis instrumen penilaian dan lebih bervariasi mengingat instrumen penilaian yang selama ini hanya menggunakan soal pilihan ganda biasa. Selain itu, belum banyaknya instrumen penilaian keterampilan berpikir kritis yang berkonteks materi kimia terutama

larutan penyangga sehingga penelitian dan pengembangan ini perlu dan penting untuk dilakukan.

Setelah perencanaan sudah matang kemudian dilakukan pengembangan produk awal. Penyusunan draf awal instrumen adalah menentukan tujuan dari penilaian. Setelah menentukan tujuan dari penilaian, maka dibuat kisi-kisinya. Kisi-kisi disusun berdasarkan kompetensi apa saja yang harus dinilai. Penelitian dan pengembangan ini bertujuan untuk mengembangkan instrumen penilaian untuk mengukur keterampilan berpikir kritis siswa. Maka oleh itu, indikator yang dibuat tidak hanya berdasar pada indikator kompetensi dari materi larutan penyangga, tetapi juga dihubungkan dengan indikator dari keterampilan berpikir kritis. Indikator keterampilan berpikir kritis yang digunakan ada enam, yaitu interpersasi, analisis, kesimpulan, evaluasi, penjelasan, dan pengaturan diri.

Draf produk yang telah selesai dibuat kemudian divalidasi. Tujuan dari dilakukannya validasi ini adalah untuk mengetahui apakah instrumen tersebut mempunyai kualitas yang baik. Kualitas yang baik ditentukan oleh isi dari instrumen tersebut yaitu meliputi kesesuaian indikator keterampilan berpikir kritis dan indikator soal dengan soal, penggunaan bahasa yang mudah dimengerti, tidak adanya kesalahan penulisan dan tanda baca, penentuan kunci jawaban yang benar. Valid tidaknya instrumen ditentukan dengan membandingkan nilai validitas hasil perhitungan dengan nilai validitas standar (V tabel). Nilai validitas dihitung menggunakan rumus Aiken.

Semua soal memiliki nilai V hitung lebih dari nilai V tabel sehingga kesimpulannya adalah semua soal valid. Walaupun semua soal bersifat valid, namun harus tetap diperbaiki sesuai dengan saran yang diberikan oleh validator agar soal lebih baik lagi baru

kemudian dilanjutkan ke uji coba selanjutnya.

Instrumen yang telah diperbaiki sesuai dengan saran dari validator dan diperoleh produk awal kemudian diujicobakan. Uji coba instrumen penilaian *Three-Tier MC* ini terdiri dari tiga tahap dengan subjek uji coba yang berbeda di setiap tahapnya.

Tahap pertama adalah uji coba lapangan awal. Uji coba ini dilakukan dengan memberikan produk awal dari instrumen penilaian *Three-Tier MC* pada 9 siswa dari setiap sekolah penelitian yang dipilih secara acak. Data dari uji coba lapangan awal ini berupa angket tanggapan dari siswa mengenai keterbacaan instrumen penilaian *Three-Tier MC*. Kesalahan pengetikan terjadi pada butir soal nomor 5 tier ke-3, yaitu $K_a \text{HNO}_3$ yang seharusnya $K_a \text{HNO}_2$.

Setelah produk awal direvisi sesuai hasil uji coba lapangan awal kemudian dilakukan tahap berikutnya yaitu uji coba lapangan utama. Produk utama hasil revisi produk awal diberikan untuk dikerjakan kepada siswa sebanyak satu kelas setiap sekolah dengan jumlah total 87 siswa. Pada tahap ini dilakukan analisis butir soal dari jawaban siswa dengan menggunakan *software* ITEMAN dimana analisis yang dilakukan meliputi reliabilitas, daya pembeda, tingkat kesukaran soal, dan indeks pengecoh.

Data hasil analisis menunjukkan bahwa tidak ada soal yang memiliki daya pembeda dengan nilai negatif. Ini berarti bahwa semua soal mampu membedakan antara kelompok bawah dengan kelompok atas. Dilihat dari tingkat kesukaran soal, proporsi soal dengan tingkat kesukaran sedang lebih mendominasi serta soal dengan tingkat kesukaran mudah dan sukar memiliki proporsi yang seimbang. Tidak adanya soal yang memiliki nilai indeks pengecoh 0,000 menunjukkan bahwa pengecoh sudah berfungsi dengan baik. Berdasarkan data analisis yang sudah

dijelaskan di atas, maka pada tahap ini tidak ada soal yang direvisi.

Tahap selanjutnya adalah uji coba pelaksanaan lapangan atau uji coba operasional. Uji coba ini pada dasarnya sama seperti uji coba lapangan utama. Produk operasional yang merupakan hasil revisi produk utama diberikan kepada siswa sebanyak dua kelas setiap sekolah dengan total jumlah 175 siswa. Sama seperti uji coba lapangan utama, jawaban siswa kemudian dianalisis menggunakan *software* ITEMAN untuk mengetahui reliabilitas, daya pembeda, tingkat kesukaran soal, dan indeks pengecoh dari instrumen penilaian *Three-Tier MC* dari setiap butir soal. Selain itu, dilakukan pula analisis keterampilan berpikir kritis siswa.

Pada soal nomor 2 *tier* 3, nomor 6 *tier* 1, nomor 8 *tier* 1, dan nomor 9 *tier* 3, mengalami penurunan daya pembeda menjadi kategori jelek. Sehingga pada soal tersebut dilakukan revisi untuk menyempurnakan produk akhir dari instrumen penilaian *Three-Tier MC* untuk mengukur keterampilan berpikir kritis siswa pada materi larutan penyangga.

Instrumen penilaian harus memiliki kualitas yang baik. Instrumen penilaian dapat dikatakan memiliki kualitas yang baik apabila memenuhi persyaratan yang tercantum pada Permendiknas Nomor 20 Tahun 2007 tentang Standar Penilaian. Menurut Sukiman agar suatu instrumen penilaian memenuhi persyaratan tersebut dapat dilakukan dua cara, yaitu 1) analisis soal secara teoritik atau kualitatif dan 2) analisis soal secara empiris atau kuantitatif (Sukiman, 2012).

Cara pertama adalah analisis secara teoritik yang dilakukan sebelum pelaksanaan uji coba. Cara ini untuk menganalisis instrumen penilaian dilihat dari segi substansi, konstruksi dan bahasa. Analisis dilakukan dengan validasi terhadap para ahli. Hasil dari validasi isi oleh para ahli dihitung

menggunakan formula Aiken. Validator berjumlah 8 orang para ahli yang terdiri dari 5 orang dosen dan 3 orang guru serta terdapat 4 skala pengukuran. Tiap soal dikatakan valid apabila memiliki nilai validitas hitung lebih dari nilai validitas tabel. Berdasarkan jumlah validator yaitu 8 orang dan skala pengukuran yaitu 4, maka nilai validitas tabelnya adalah 0,75. Hasil perhitungan validitas terhadap 15 soal menunjukkan nilai validitasnya lebih dari 0,75 sehingga dapat dikatakan bahwa semua soal sudah valid. Akan tetapi perbaikan soal juga dilakukan sesuai dengan saran dari masing-masing validator agar soal yang telah dibuat menjadi lebih baik lagi.

Cara kedua adalah analisis secara empiris. Cara empiris dilakukan dengan cara uji coba untuk mengetahui reliabilitas, daya pembeda, tingkat kesukaran, dan indeks pengecoh soal. Uji coba terdiri dari 2 kali pelaksanaan. Uji coba pertama dengan mengikutsertakan siswa sebanyak 87 siswa dan uji coba kedua sebanyak 175 siswa sehingga jumlah totalnya adalah 262 siswa. Siswa yang menjadi subjek penelitian ini berasal dari 3 sekolah yang terdiri dari sekolah kategori tinggi, sedang, dan rendah. Hal ini dilakukan agar instrumen penilaian yang dikembangkan dapat diterapkan pada siswa dengan kemampuan rendah sampai tinggi. Dari hasil analisis empiris, soal-soal tersebut telah dapat dikatakan cukup baik dilihat dari reliabilitas, daya pembeda, tingkat kesukaran soal dan indeks pengecohnya. Dimana reliabilitasnya termasuk dalam reliabilitas sedang, daya pembedanya tidak ada yang bernilai negatif, soal yang memiliki tingkat kesukaran mudah dan sukar tidak mendominasi, dan semua pengecoh berfungsi dengan baik yang ditunjukkan tidak adanya nilai indeks pengecoh 0,000.

Instrumen penilaian pada umumnya dibuat hanya untuk mengukur kemampuan kognitif siswa saja. Penelitian dan pengembangan ini

bertujuan untuk mengembangkan instrumen penilaian yang tidak hanya untuk mengukur kemampuan kognitif siswa akan tetapi dapat digunakan untuk mengukur keterampilan berpikir kritis siswa. Keterampilan berpikir kritis di dalam penelitian dan pengembangan instrumen penilaian ini mengacu pada keterampilan berpikir kritis yang dikemukakan oleh Facione dengan enam indikator yaitu interpretasi, analisis, evaluasi, inferensi, penjelasan dan pengaturan diri. Keenam indikator tersebut adalah inti dari keterampilan berpikir kritis yang menurut para ahli termasuk didalam keterampilan kognitif (Facione, 2015).

Instrumen penilain yang dikembangkan berupa soal pilihan ganda bertingkat yaitu *Three-Tier Multiple Choice*. Soal pilihan ganda dipilih karena lebih efisien dalam pengoreksiannya dan objektif. Kekurangannya adalah pada soal bentuk pilihan ganda sulit untuk mengetahui proses berpikir siswa dalam mengerjakan soal tersebut. Maka soal pilihan ganda ini dikembangkan menjadi pilihan ganda bertingkat. Tujuannya agar dapat mengetahui bagaimana proses berpikir siswa dalam menjawab soal atau memecahkan masalah. Tingkat pertama merupakan soal pengetahuan yang mewakili beberapa indikator dari keterampilan berpikir kritis seperti interpretasi, analisis, evaluasi, dan kesimpulan. Tingkat kedua merupakan pilihan alasan jawaban dari soal tingkat pertama sehingga siswa harus mengetahui penjelasan dari tingkat pertama. Sedangkan tingkat ketiga mewakili indikator *self regulation* (pengaturan diri).

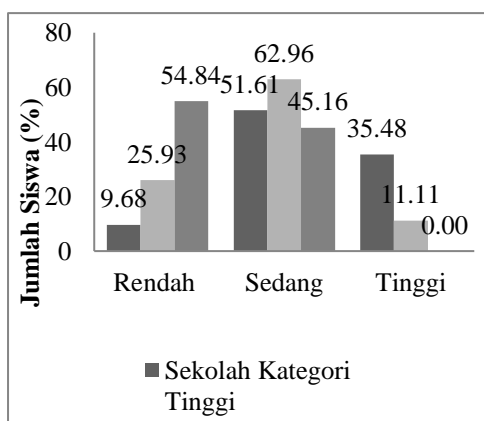
Soal tes bertingkat (dua tingkat maupun tiga tingkat) yang telah dikembangkan merupakan tes *diagnostic* untuk mengetahui miskonsepsi yang terjadi. Seperti yang dilakukan oleh Pesman, dia melakukan pengembangan tes tiga tingkat sebagai instrumen diagnostik untuk mengetahui

miskonsepsi pada Sirkuit Elektrik Sederhana (Peşman & Eryilmaz, 2010). Pengembangan tes tiga tingkat juga dilakukan oleh Eryilmaz untuk menilai miskonsepsi tentang Panas dan Temperatur (Eryilmaz, 2010). Sebelumnya, pengembangan instrumen diagnostik dua tingkat telah dilakukan oleh Tuysuz pada materi Pemisahan Zat. Berbeda dengan tes bertingkat yang telah dikembangkan oleh para peneliti sebelumnya, instrumen penilaian *Three-Tier MC* yang peneliti kembangkan ini merupakan tes *placement*, untuk mengetahui tingkat keterampilan berpikir kritis siswa pada materi Larutan Penyangga. Menurut Kartimi berpikir kritis termasuk karakteristik psikologis seseorang yang dapat diketahui kualifikasinya (rendah, sedang, atau tinggi) dan hal itu bisa diketahui apabila diadakan pengukuran dengan aturan dan formula yang jelas (Kartimi & Liliarsari, 2012).

Hasil perbandingan persentase keterampilan berpikir kritis siswa pada Gambar 1 menunjukkan keterkaitan antara keterampilan berpikir kritis siswa dengan kemampuan atau hasil belajar siswa. Sekolah yang menjadi subjek penelitian memiliki tiga kategori kemampuan yang berbeda, yaitu sekolah dengan kemampuan kategori tinggi, sedang, dan rendah. Hal ini sesuai dengan hasil yang diperoleh dari analisis keterampilan berpikir kritis siswa bahwa sekolah kategori tinggi memiliki keterampilan berpikir kritis paling baik dan sekolah kategori rendah memiliki keterampilan berpikir kritis paling buruk. Hal ini juga sesuai dengan pendapat yang dikemukakan oleh Qing bahwa di dalam proses belajar diperlukan kemampuan berpikir (Qing, Jing, Yazhuan, Ting, & Junping, 2010). Ketika siswa-siswinya memiliki keterampilan berpikir kritis yang baik atau tinggi maka hasil belajarnya juga akan baik atau tinggi.

Hasil ini juga sesuai dengan penelitian yang telah dilakukan oleh

Kartimi yaitu mengembangkan alat ukur berpikir kritis, dimana hasilnya menunjukkan bahwa alat ukur yang dikembangkan mampu membedakan keterampilan berpikir kritis siswa (Kartimi & Liliarsari, 2012). Kartimi mengembangkan alat ukur berpikir kritis pada konsep hidrokarbon, termokimia, dan kesetimbangan kimia. Bentuk soalnya juga berupa pilihan ganda bertingkat tetapi hanya terdiri dari dua tingkat. Indikator keterampilan berpikir kritis yang digunakan adalah lima indikator yang dikembangkan oleh Ennis.



Gambar 1. Perbandingan Keterampilan Berpikir Kritis Siswa

Penelitian yang dilakukan oleh Zoller (2007) untuk mengembangkan keterampilan kognitif tingkat tinggi siswa atau *higher-order cognitive skills* (HOCS) pada kimia organik menyatakan bahwa HOCS dalam pendidikan sains berkaitan erat dengan berpikir kritis dan pemecahan masalah. Ketika siswa memiliki keterampilan berpikir kritis yang baik maka keterampilan kognitifnya juga akan baik. Hasil dari penelitian yang dilakukan oleh Miri menyimpulkan apabila guru secara terus menerus melatih siswa dengan menggunakan strategi pembelajaran untuk meningkatkan berpikir tingkat tinggi ada kesempatan yang baik untuk mengembangkan keterampilan berpikir kritis (Miri, David, & Uri, 2007). Mengerjakan soal-soal merupakan salah

satu bentuk latihan siswa untuk mengasah keterampilan siswa. Apabila siswa sering berlatih soal menggunakan instrumen penilaian *Three-Tier MC* maka diharapkan hasil belajar dan keterampilan berpikir kritis siswa akan meningkat. Salah satu soal yang terdapat pada instrumen penilaian *Three-Tier MC* adalah sebagai berikut:

Diantara campuran berikut, yang dapat membentuk larutan penyangga adalah

- A. HCl dan NH_4Cl
- B. NaOH dan CH_3COONa
- C. HBr dan NaBr
- D. KOH dan KCl
- E. NH_4OH dan NH_4Cl

Alasan jawaban

- A. Larutan penyangga merupakan campuran antara asam kuat dengan asam konjugasinya
- B. Larutan penyangga merupakan campuran antara basa kuat dengan basa konjugasinya
- C. Larutan penyangga merupakan campuran antara asam lemah dengan basa konjugasinya
- D. Larutan penyangga merupakan campuran antara basa lemah dengan asam konjugasinya

NH_3 dan NH_4Br merupakan campuran larutan penyangga.

- A. Benar
- B. Salah

Penelitian dan pengembangan instrumen penilaian ini dapat digunakan sebagai alternatif penilaian untuk mengukur keterampilan berpikir kritis siswa. Instrumen penilaian *Three-Tier MC* memiliki tiga fungsi sekaligus. Pertama, fungsi formatif yaitu untuk menentukan umpan balik (*feedback*) baik bagi guru maupun siswa, digunakan sebagai dasar untuk memperbaiki proses pembelajaran dan mengadakan program *remedial* bagi siswa. Kedua, fungsi

penempatan yaitu mengetahui tingkat keterampilan berpikir kritis dari masing-masing siswa, karena pada dasarnya siswa memiliki kemampuan yang berbeda serta kelemahan dan kelebihan masing-masing. Siswa memiliki keterampilan berpikir kritis yang berbeda-beda dari rendah sampai tinggi. Sehingga guru dapat menempatkan siswa dalam situasi pembelajaran yang tepat sesuai dengan tingkat kemampuan siswa. Ketiga, fungsi diagnostik yaitu untuk mengetahui kelemahan siswa beserta penyebabnya. Bagian dari bahan pelajaran yang masih sulit dikuasai dapat diketahui dengan instrumen penilaian ini. Tiga fungsi tersebut dapat digunakan guru untuk mengetahui sejauh mana materi pelajaran dapat diterima siswa, apakah tujuan pembelajaran sudah tercapai atau belum, serta membantu guru untuk menentukan langkah yang tepat untuk memperbaikinya.

Kesimpulan dan Rekomendasi

Berdasarkan penelitian dan pengembangan yang telah dilakukan, maka dapat ditarik simpulan bahwa instrumen penilaian *Three-Tier MC* yang telah dikembangkan termasuk dalam kategori yang baik ditinjau dari segi validitasnya yang lebih dari 0,75 dan dikatakan valid; reliabilitasnya termasuk dalam sedang, daya pembedanya tidak ada yang bernilai negatif dengan persentase jelek 8,89%, cukup baik 20%, baik 24,44%, sangat baik 46,67%; soal yang memiliki tingkat kesukaran mudah dan sukar tidak mendominasi dengan persentase soal mudah 15,56%, sedang 73,33%, sukar 11,11%; dan semua pengecoh berfungsi dengan baik yang ditunjukkan tidak adanya nilai indeks pengecoh 0,000; serta instrumen penilaian *Three-Tier MC* yang telah dikembangkan pada materi larutan penyangga dapat digunakan untuk mengukur keterampilan berpikir kritis siswa.

Daftar Pustaka

- Bayrak, B. K. (2013). Using Two-Tier Test to Identify Primary Students' Conceptual Understanding and Alternative Conceptions in Acid Base. *Mevlana International Journal of Education*, 3(2), 19–26.
- Borg, W. R., & Gall, M. D. (1983). *Educational Research: An Introduction* (4th ed.). New York: Longman.
- Chandrasegaran, a. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice*, 8(3), 293.
- Eryilmaz, A. (2010). Development and Application of Three-Tier Heat and Temperature Test: Sample of Bachelor and Graduate Students. *Eurasian Journal of Educational Research*, (40), 53–76.
- Facione, P. A. (2015). *Critical Thinking: What It Is and Why It Counts. Insight Assessment*.
- Fisher, A. (2009). *Berpikir Kritis*. Jakarta: Erlangga.
- Kartimi, & Liliarsari. (2012). Pengembangan Alat Ukur Berpikir kritis pada Konsep Termokimia untuk Siswa SMA Peringkat Atas dan Menengah. *Jurnal Pendidikan IPA Indonesia (JPPI)*, 1(1), 21–26.
- Kartimi, Liliarsari, & Permanasari, A. (2012). Pengembangan Alat Ukur Berpikir Kritis pada Konsep Senyawa Hidrokarbon untuk Siswa SMA di Kabupaten Kuningan. *Jurnal Pendidikan MIPA*, 13(1), 18–25.
- Miri, B., David, B., & Uri, Z. (2007). Purposely Teaching for the Promotion of Higher-order Thinking Skills: A Case of Critical Thinking. *Res Sci Educ*, 37, 353–369.
- Nofiana, M., Sajidan, & Puguh. (2014).

- Pengembangan Instrumen Evaluasi Two-Tier Multiple Choice Question untuk Mengukur Keterampilan Berpikir Tingkat Tinggi. *Jurnal Inkuiri*, 3(II), 60–74.
- Peşman, H., & Eryilmaz, A. (2010). Development of a Three-Tier Test to Assess Misconceptions About Simple Electric Circuits. *The Journal of Educational Research*, 103(3), 208–222.
- Prasetyowati, E. N., & Suyatno. (2016). Peningkatan Penguasaan Konsep dan Keterampilan Berpikir Kritis Siswa Melalui Implementasi Model Pembelajaran Inkuiri pada Materi Pokok Larutan Penyangga. *Jurnal Kimia Dan Pendidikan Kimia (JKPK)*, 1(1), 67–74.
- Qing, Z., Jing, G., Yazhuan, L., Ting, W., & Junping, M. (2010). Promoting Preservice Teacher's Critical Thinking Disposition by Inquiry-Based Chemical Experiment. *Procedia - Social and Behavioral Sciences*, 9, 1429–1436.
- Qing, Z., Ni, S., & Hong, T. (2010). Developing Critical Thinking Disposition by Task-Based Learning in Chemistry Experiment Teaching. *Procedia - Social and Behavioral Sciences*, 2(2), 4561–4570.
- Sukiman. (2012). Pengembangan Sistem Evaluasi. Yogyakarta: Insan Mandiri.
- Sulfivan, M. J., & Dawson, C. J. (1979). Critical Thinking Ability in Science - What can be Measured? *Research in Science Education*, 9, 159–167.
- Yamtinah, S., Saputro, S., & Utami, B. (2016). Content Validity and Scoring of Two Tier as Measuring Instrument of Science Process Skills for Knowledge Aspects in Chemistry Learning. *Prosiding ICTTE FKIP UNS 2015*, 1(1), 911–916.